microRNA Profiling Identifies Cancer-Specific and Prognostic Signatures in Pediatric Malignancies

Jun S. Wei,[1] Peter Johansson,[1] Qing-Rong Chen,[1,2] Young K. Song,[1] Steffen Durinck,[1] Xinyu Wen,[1,2] Adam T.C. Cheuk,[1] Malcolm A. Smith,[3] Peter Houghton,[4] Christopher Morton,[4] and Javed Khan[1]

Abstract

Purpose: microRNAs have been shown to be involved in different human cancers. We therefore have performed expression profiles on a panel of pediatric tumors to identify cancer-specific microRNAs. We also investigated if microRNAs are coregulated with their host gene.

Experimental Design: We performed parallel microRNAs and mRNA expression profiling on 57 tumor xenografts and cell lines representing 10 different pediatric solid tumors using microarrays. For those microRNAs that map to their host mRNA, we calculated correlations between them.

Results: We found that the majority of cancer types clustered together based on their global microRNA expression profiles by unsupervised hierarchical clustering. Fourteen microRNAs were significantly differentially expressed between rhabdomyosarcoma and neuroblastoma, and 8 of them were validated in independent patient tumor samples. Exploration of the expression of microRNAs in relationship with their host genes showed that the expression for 43 of 68 (63%) microRNAs located inside known coding genes was significantly correlated with that of their host genes. Among these 43 microRNAs, 5 of 7 microRNAs in the OncomiR-1 cluster correlated significantly with their host gene *MIRHG1* ($P < 0.01$). In addition, high expression of *MIRHG1* was significantly associated with high stage and *MYCN* amplification in neuroblastoma tumors, and the expression level of *MIRHG1* could predict the outcome of neuroblastoma patients independently from the current neuroblastoma risk-stratification in two independent patient cohorts.

Conclusion: Pediatric cancers express cancer-specific microRNAs. The high expression of the OncomiR-1 host gene *MIRHG1* correlates with poor outcome for patients with neuroblastoma, indicating important oncogenic functions of this microRNA cluster in neuroblastoma biology. (Clin Cancer Res 2009;15(17):5560–8)

MicroRNAs are small, noncoding RNA molecules encoded in the genomes of plants and animals. These highly conserved, ~21-nucleotide RNAs regulate the expression of genes by binding to the 3′-untranslated regions of specific mRNAs, causing translational inhibition or mRNA degradation (1). As many mRNAs may share this short sequence, microRNAs are capable of simultaneously influencing the expression of large sets of genes. It is estimated that each microRNA can target hundreds of genes (2); conversely, multiple microRNAs can target a single gene. Thus far, 701 microRNAs (version 12.0) have been reported to be expressed in human cells.[5] Due to their regulatory roles in gene expression, there is increasing evidence that microRNAs are directly involved not only in normal embryogenesis,

[5] http://microrna.sanger.ac.uk/

Authors' Affiliations: [1]Oncogenomics Section, Pediatric Oncology Branch, Advanced Technology Center, National Cancer Institute, Gaithersburg, Maryland; [2]Advanced Biomedical Computing Center, Science Applications International Corporation-Frederick, Inc., National Cancer Institute-Frederick, Frederick, Maryland; [3]Cancer Therapy Evaluation Program, National Cancer Institute, Bethesda, Maryland; and [4]St. Jude Children's Research Hospital, Memphis, Tennessee

Received 12/19/08; revised 6/24/09; accepted 7/20/09; published OnlineFirst 8/25/09.

Grant support: Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

J.S. Wei and P. Johansson contributed equally to this work.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

Requests for reprints: Jun S. Wei, Oncogenomics Section, Pediatric Oncology Branch, Advanced Technology Center, National Cancer Institute, 8717 Grovemont Circle, Room 225D, Gaithersburg, MD 20877. Phone: 301-402-9031; Fax: 301-402-3134; E-mail: weij@mail.nih.gov and Javed Khan, Oncogenomics Section, Pediatric Oncology Branch, Advanced Technology Center, National Cancer Institute, 8717 Grovemont Circle, Room 134E, Gaithersburg, MD 20877. Phone: 301-435-2937; Fax: 301-480-0314; E-mail: khanjav@mail.nih.gov.

doi:10.1158/1078-0432.CCR-08-3287

Translational Relevance

microRNAs are small, noncoding regulatory RNAs that are implicated in cancer development. Because most pediatric malignancies are developmental tumors arising from aberrant differentiation, we hypothesized that pediatric tumors will exhibit cancer and tissue-specific microRNA expression profiles associating with development and tumorigenic process, which can be used in diagnosis and prognosis. Here, we have performed microRNAs and mRNA expression profiling on a panel of 57 tumor xenografts and cell lines representing 10 different pediatric solid tumors using microarrays. We showed that pediatric cancers differentially express microRNAs specific to their origins and types. In addition, we showed evidence that the expression of microRNAs located within protein coding genes is coregulated with their host gene transcripts. Finally, we showed that the high expression of the OncomiR-1 host gene, *MIRHG1*, is significantly associated with aggressive neuroblastoma. This finding warrants further studies of the role of the OncomiR-1 in neuroblastoma patients with adverse outcomes.

metabolism, cell growth, differentiation, and apoptosis but also in pathogenesis of human cancers (3–6). Because most pediatric malignancies are developmental tumors arising from aberrant differentiation, we hypothesized that pediatric tumors will exhibit cancer and tissue-specific microRNA expression profiles associating with development and the tumorigenic process, which can be used in classification and prognosis of cancers.

To test this hypothesis, we investigated the expression profiles of microRNAs for a panel of 57 pediatric cell lines and human tumor xenografts for which mRNA profiles were available, and the majority of which are currently used as pediatric preclinical models for drug screening (7, 8). Using this panel of samples representing 10 different types of pediatric tumors (Table 1), we explored whether pediatric tumors differentially express microRNAs according to their diagnosis using microarray technology. A machine learning algorithm and statistical analysis was applied to the microRNA expression data to identify tumor-specific profiles for the two major subgroups of cancers (neuroblastoma and rhabdomyosarcoma) represented in our data set, and we validated these findings on independent neuroblastoma and rhabdomyosarcoma patient tumor samples. We explored if the microRNAs that map within host messenger RNAs are coregulated with their host mRNAs. Finally, we investigated if the expression of *MIRHG1* gene (formally *C13orf25*), which hosts the oncogenic microRNA OncomiR-1 (miR-17-92 cluster), correlated with aggressive disease and poor outcome for patients with neuroblastoma.

Materials and Methods

Cell lines, xenografts, and primary tumor samples. Neuroblastoma cell lines (n = 16) were cultured as described (9). Xenograft samples (n = 41) were described elsewhere (7, 8) and obtained through the Pediatric Preclinical Testing Program established by the National Cancer Institute (NCI). Anonymous primary snap-frozen neuroblastoma (n = 6) and rhabdomyosarcoma tumors (n = 6) were acquired from Corporative Human Tissue Network and were deemed exempt from NCI institutional review board for this study. The clinical characteristics of these primary samples are described in Supplementary Table S1.

Microarray and TaqMan real-time reverse transcription-PCR assays for microRNA expression. Small RNA (<200 bp) was purified using a previously published protocol (10). microRNA expression profiling was done on our in-house printed microarrays. Synthetic DNA probes were designed using the sequences available from the Sanger miRBase

Table 1. Summary of samples

Sample name	Type	Diagnosis
D212-X	Xenograft	Glioblastoma
D456-X	Xenograft	Glioblastoma
SJBT39-X	Xenograft	Glioblastoma
SJBT56-X	Xenograft	Glioblastoma
SJGBM2-X	Xenograft	Glioblastoma
BT45-X	Xenograft	BT other (medulloblastoma)
BT46-X	Xenograft	BT other (medulloblastoma)
BT50-X	Xenograft	BT other (medulloblastoma)
BT36-X	Xenograft	BT other (ependymoma)
BT41-X	Xenograft	BT other (ependymoma)
ASLuc-C	Cell line	Neuroblastoma
BE2-C	Cell line	Neuroblastoma
CHP134-C	Cell line	Neuroblastoma
GILIN-C	Cell line	Neuroblastoma
IMR32-C	Cell line	Neuroblastoma
IMR5-C	Cell line	Neuroblastoma
KCNR-C	Cell line	Neuroblastoma
LAN1-C	Cell line	Neuroblastoma
LAN5-C	Cell line	Neuroblastoma
NB1691-C	Cell line	Neuroblastoma
NBEB-C	Cell line	Neuroblastoma
SKNAS-C	Cell line	Neuroblastoma
SKNDZ-C	Cell line	Neuroblastoma
SKNFI-C	Cell line	Neuroblastoma
SKNSH-C	Cell line	Neuroblastoma
SY5Y-C	Cell line	Neuroblastoma
CHLA79-X	Xenograft	Neuroblastoma
NB-1382-X	Xenograft	Neuroblastoma
NB1643-X	Xenograft	Neuroblastoma
NB1691-X	Xenograft	Neuroblastoma
NB1771-X	Xenograft	Neuroblastoma
NBEBc1-X	Xenograft	Neuroblastoma
NBSD-X	Xenograft	Neuroblastoma
SKNAS-X	Xenograft	Neuroblastoma
OS1-X	Xenograft	Osteosarcoma
OS17-X	Xenograft	Osteosarcoma
OS2-X	Xenograft	Osteosarcoma
OS21-X	Xenograft	Osteosarcoma
SKNEP-X	Xenograft	Other (diffuse anaplastic Wilms' tumor)
EW5-X	Xenograft	Other (Ewing's)
EW8-X	Xenograft	Other (Ewing's)
BT29-X	Xenograft	Other (rhabdoid tumor of brain)
KT12-X	Xenograft	Other (rhabdoid tumor of kidney)
KT16-X	Xenograft	Other (rhabdoid tumor of kidney)
Unknown1-X	Xenograft	Other (unknown)
Unknown2-X	Xenograft	Other (unknown)
Rh28-X	Xenograft	Rhabdomyosarcoma
Rh30-X	Xenograft	Rhabdomyosarcoma
RH30R-X	Xenograft	Rhabdomyosarcoma
Rh36-X	Xenograft	Rhabdomyosarcoma
Rh41-X	Xenograft	Rhabdomyosarcoma
Rh65-X	Xenograft	Rhabdomyosarcoma
KT10-X	Xenograft	Wilms' tumor
KT11-X	Xenograft	Wilms' tumor
KT13-X	Xenograft	Wilms' tumor

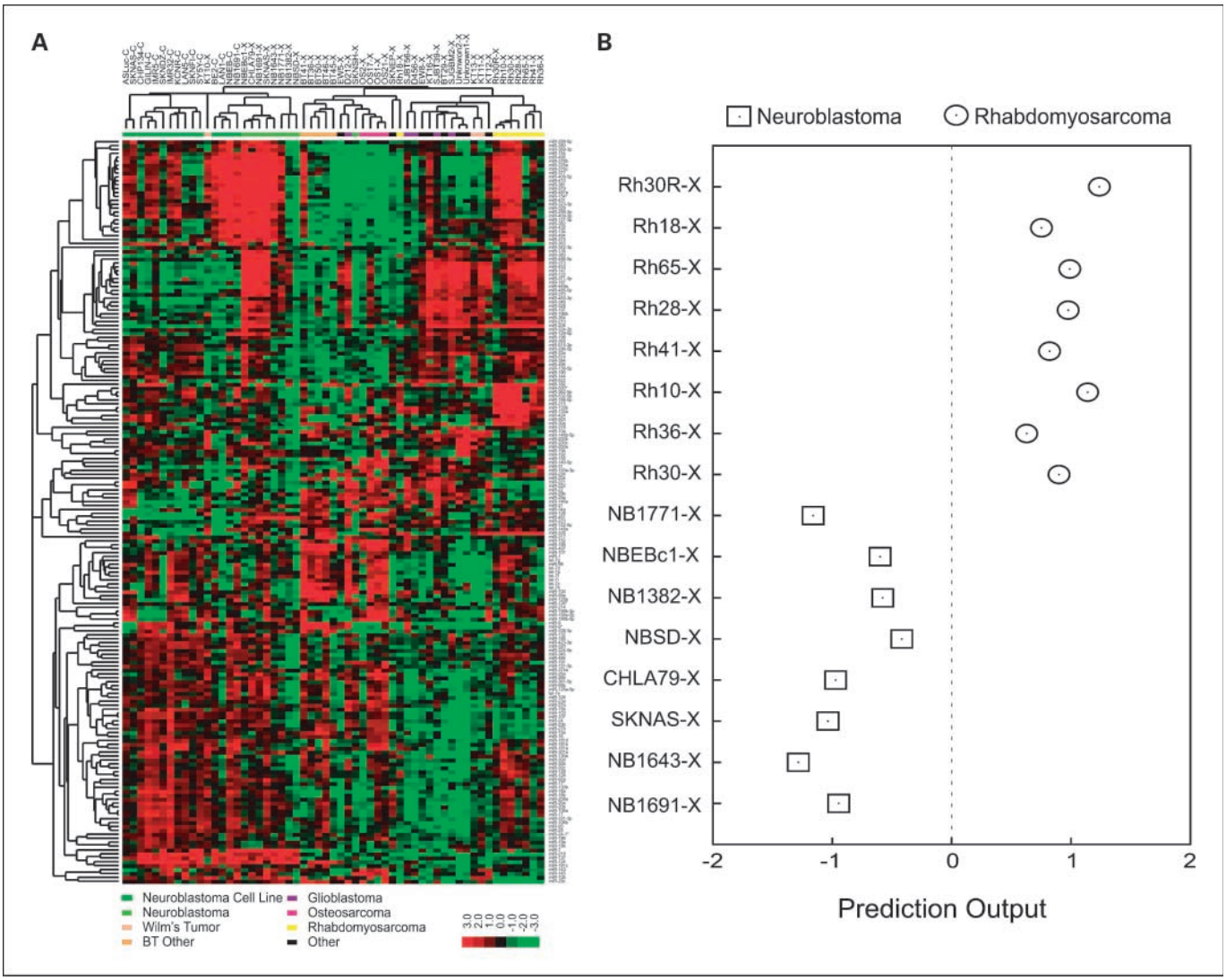

Fig. 1. microRNAs are diagnostic for pediatric tumor samples. *A,* unsupervised hierarchical clustering showed clustering of the samples according to the tumor types. We performed two-way hierarchical clustering of 57 pediatric cancer samples using all microRNA expression values after quality filtering. The Cluster 3.0 software (12) was used, wherein expression values were median-centered per gene and clustered using Pearson correlation distance and average linkage. *B,* NCCs were trained to separate neuroblastoma from rhabdomyosarcoma xenograft samples. The centroid for each diagnostic category was calculated as the average profile across samples. The prediction output of a sample was calculated as the Pearson correlation against the rhabdomyosarcoma centroid minus the Pearson correlation against the neuroblastoma centroid. Prediction accuracy was evaluated using a leave-one-out scheme, and all samples were correctly classified ($P = 1.6 \times 10^{-4}$).

Sequence Database (11) and custom-made by Sigma. Each probe contained two tandem complementary sequences against each mature microRNA or its counterpart strand in the hairpin stem-loop structure, and there were 521 unique probes for human microRNAs on our in-house microRNA microarrays. An amine group tag was added on the 5′ of each probe for tethering on Nexterion epoxy glass slides (Schott). The reference synthetic DNA oligos complementary to the probes were labeled with Cy3 dye and the sample with Cy5 dye using miRVana microRNA labeling kit (Ambion). Hybridization was done on the MAUI hybridization systems (BioMicro System) at 52°C with mixing for overnight. Slides were then washed in 2× SSC with 0.2% SDS for 15 min at 42°C, 2× SSC at room temperature for 10 min, and 0.2× SSC at room temperature for 10 min. Finally, slides were dried by centrifugation and scanned in an Agilent microarray scanner (Agilent).

TaqMan microRNA reverse transcription-PCR assays (Applied Biosystems) were done according to the manufacturer's protocol as described previously (10).

microRNA microarray data filtering and normalization. The Cy3 reference channel was first normalized using quantile normalization. Then, low-quality probes were removed using the Cy5 sample channel with a criterion that required raw intensity of the probe to be larger than 128 fluorescent units for at least 4 samples. Two hundred seven probes passed this quality filter. After quality filtering, the $\log_2$ ratios [$\log_2$(Cy5/Cy3)] were calculated and subsequently normalized by subtracting the average $\log_2$ ratio of the internal control probes. Then, we added a constant value to get positive values. Every probe was printed in duplicate on the array and the average of these duplicates was used to represent the final expression measurements. All of the quality-filtered microRNA and parallel mRNA data can be found on our Web site.[6]

mRNA microarray experiments. Gene expression profiling was done on Affymetrix U133 Plus 2.0 arrays according to the manufacturer's instruction (Affymetrix). We obtained the gene expression profiling data

[6] http://pob.abcc.ncifcrf.gov/cgi-bin/JK

of 38 xenografts from the study by Neale et al. (7) and profiled the rest samples in our laboratory. Relative expression values were obtained using Affymetrix PLIER algorithm through Affymetrix Power Tools version 1.8.5 and further log-transformed to base 2.

Hierarchical clustering, nearest centroid classifiers, and statistical and genomic location analyses. We performed two-way hierarchical clustering of 57 pediatric cancer samples using all expression values from microRNA after quality filtering. The Cluster 3.0 software (12) was used, wherein expression values were median-centered per gene and clustered using Pearson correlation distance and average linkage. The result was visualized using TreeView (13).

Nearest centroid classifiers (NCC; ref. 14) were trained to separate neuroblastoma from rhabdomyosarcoma samples. In the NCC, the centroid for each class was calculated as the average profile across samples. For a test sample, the prediction output was calculated as the Pearson correlation against the rhabdomyosarcoma centroid minus the Pearson correlation against the neuroblastoma centroid. Therefore, a large prediction output suggests that the sample is a rhabdomyosarcoma sample. Prediction accuracy was evaluated using a leave-one-out scheme, in which all but one samples were used for training and the status of the left out sample was predicted by the trained classifier. We used a permutation test to estimate statistical significance of prediction accuracy. Sample labels were randomly permuted 100,000 times, and for each randomization, the leave-one-out procedure was repeated, and a *P* value was calculated corresponding to the probability to obtain perfect predictions for random sample labels.

To identify the host genes for microRNAs, genomic locations for microRNAs and mRNA probe sets were retrieved from Sanger miRBase (version 10.1) and Affymetrix, respectively, and were mapped to the UniGene. The host genes (represented by probe sets) were identified for each microRNA by mapping them in the same UniGene with the same orientation. We calculated the Pearson correlation across all 57 samples. If a microRNA was matched to several probe sets, the largest absolute correlation was used.

Survival analysis. Cox regression analyses and log-rank tests were done using the survival R package.[7] Expression values of a gene were dichotomized into high and low expression using the median as cutoffs.

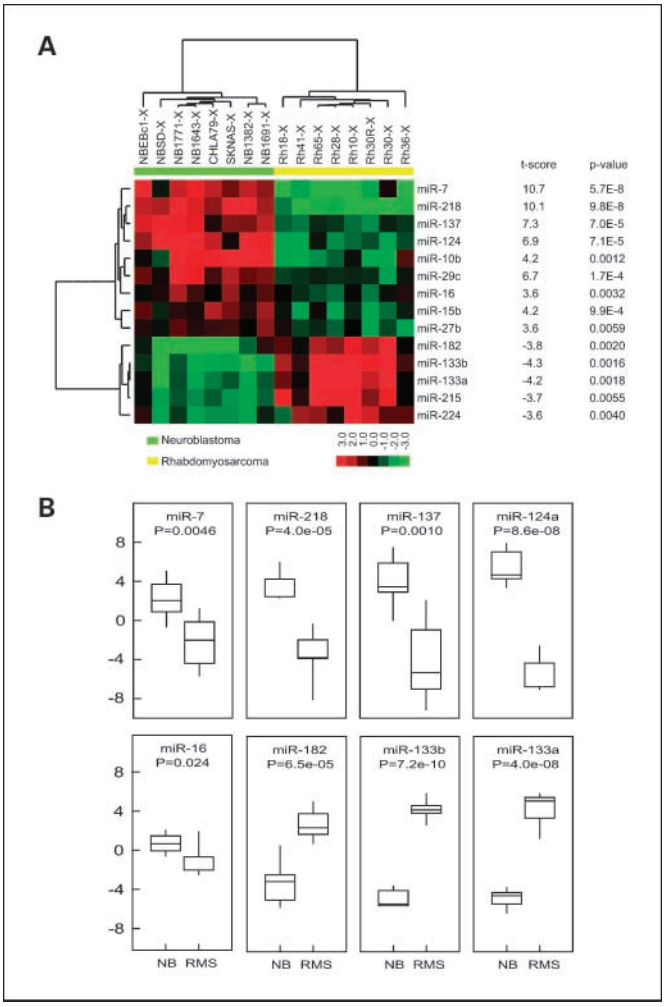

Fig. 2. Differentially expressed cancer-specific microRNAs for neuroblastoma and rhabdomyosarcoma. *A*, comparing the microRNA expression in rhabdomyosarcoma ($n = 8$) and neuroblastoma ($n = 8$) xenograft samples using a *t* test, we identified 14 differentially expressed microRNAs ($P < 0.01$). Samples and microRNAs were hierarchically clustered using Pearson correlation distance and average linkage. Data were centralized prior clustering such that median expression of each microRNA was zero. *B*, expression of 8 differentially expressed microRNAs was validated using TaqMan real-time reverse transcription-PCR in an independent set of primary neuroblastoma and rhabdomyosarcoma tumors ($P < 0.01$), indicating the value of microRNA expression levels in distinguishing these cancers.

Results

microRNA expression profiling of pediatric cancers and classification of cancers using microRNA profiles. We first hypothesized that the microRNA expression profiles would reflect the cancer type for pediatric malignancies. We performed microRNA microarray analysis of 16 neuroblastoma cell lines and 41 xenografts including brain tumor ($n = 10$), neuroblastoma ($n = 8$), rhabdomyosarcoma ($n = 8$), osteosarcoma ($n = 4$), Wilms' tumor ($n = 3$), and others ($n = 8$; Table 1). An unsupervised hierarchical clustering analysis using all 207 microRNA probes with good quality showed that microRNA expression profiles can separate these samples according to their diagnosis (Fig. 1A). One of the two major branches consists of all except one neuroblastoma samples, indicating that these samples have a neuroblastoma-specific microRNA expression profile. The other major branch contains nonneuroblastoma samples formed clusters primarily of the same diagnostic categories. For example, 7 of 8 rhabdomyosarcomas form a tight subcluster as did 4 of 4 osteosarcomas (Fig. 1A). The property of the sample clustering shows that there is a clear tumor-specific microRNA expression profile in these pediatric cancer samples.

To further examine if microRNAs can be used to classify cancers, we applied a machine learning algorithm to the microRNA expression data for the two major subgroups of cancers (neuroblastoma and rhabdomyosarcoma) represented in our data set using all high-quality probes. To avoid classifications heavily driven by cell line–specific signatures, we used only xenograft samples in this analysis. We built NCCs (14) to separate the 8 neuroblastoma xenografts from the 8 rhabdomyosarcoma xenografts using a leave-one-out scheme. All 16 samples were perfectly diagnosed ($P = 1.6 \times 10^{-4}$; Fig. 1B). Therefore, these experiments showed that neuroblastoma and rhabdomyosarcoma differentially express tumor-specific microRNAs.

Differential expression of microRNAs distinguishes neuroblastoma versus rhabdomyosarcoma. To identify the tumor-specific microRNAs that may contribute to the biology of these cancers, we performed a *t* statistical test between the two major cancer types, neuroblastoma and rhabdomyosarcoma, using the microRNA expression profiles of 8 neuroblastoma and 8 rhabdomyosarcoma xenograft samples. We found 14 microRNAs

[7] http://cran.r-project.org/web/packages/survival/index.html

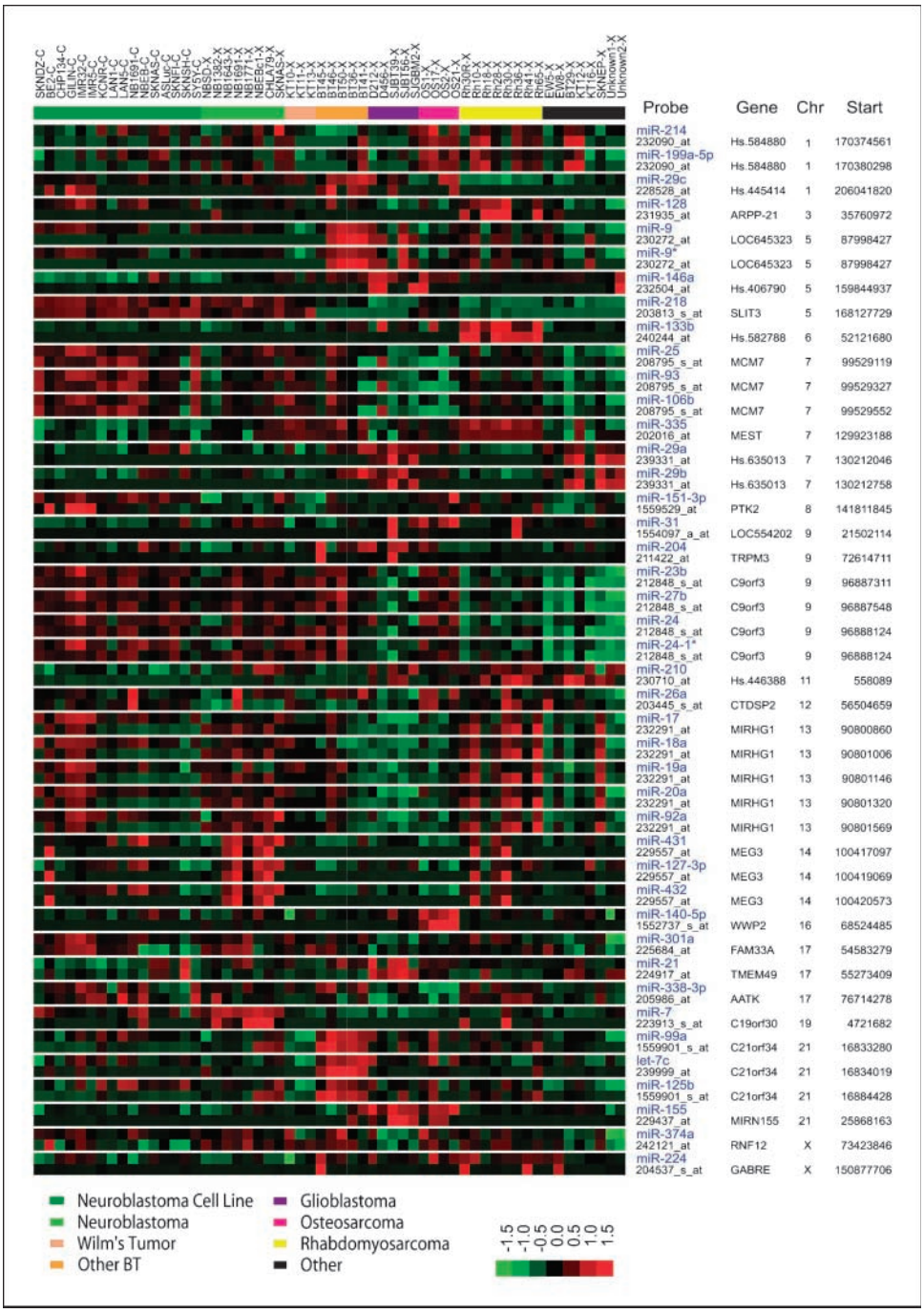

Fig. 3. microRNA expression correlates with their host gene expression. Genomic locations for microRNAs and probe sets were retrieved from Sanger miRBase (version 10.1) and Affymetrix, respectively. Genomic locations were mapped to UniGene clusters. Off 189 microRNAs with known locations, 68 (36%) were located within a UniGene with the same orientation. For each pair of microRNA and its host gene located within the same UniGene, the Pearson correlation was calculated across all 57 samples. Of 68 microRNAs located within a UniGene, 43 (63%) were found to be significantly ($P < 0.01$) positively correlated with a probe set from its host gene. Colors in the heat map represent *z*-score normalized (zero mean and unity variance per microRNA/gene) expression. Samples (*columns*) are sorted with respect to diagnosis, and mRNAs or microRNAs (*rows*) are sorted with respect to genomic location. *Chr,* chromosome; *Start,* start coordinate of microRNA.

significantly differentially expressed in these two cancer types ($P < 0.01$; false discovery rate = 15%; Fig. 2A). To validate if these microRNAs are differentially expressed in human primary tumor samples, we performed TaqMan real-time reverse transcription-PCR using an independent set of primary neuroblastoma ($n = 6$) and rhabdomyosarcoma ($n = 6$) tumors from patient biopsies. We found 8 of the 14 microRNAs to be significantly differentially expressed in the patient tumor samples ($P < 0.05$; Fig. 2B), showing the potential of using these microRNAs as biomarkers to distinguish these two cancer types.

Coregulation of microRNA with host gene. Currently, the control of microRNA expression is largely unclear, and we hypothesized that the expression level of microRNAs mapping within coding genes is controlled by the promoter of the host gene. We therefore investigated how many of the microRNAs on our array map are within and also correlate with their host mRNAs. Of 207 microRNAs of good quality detected by our arrays, 189 unique microRNAs can be mapped to the human genome with their genomic coordinates. Among them, 68 (36%) are located within known coding genes and also have the same orientation with their host genes. We calculated correlations of the expression between these 68 microRNAs and their corresponding mRNAs and found that 43 (63%) were significantly coexpressed ($r > 0.34$; $P < 0.01$) with their host genes (Fig. 3; Supplementary Table S2). OncomiR-1, also known as the miR-17-92 cluster, has been reported to be associated with

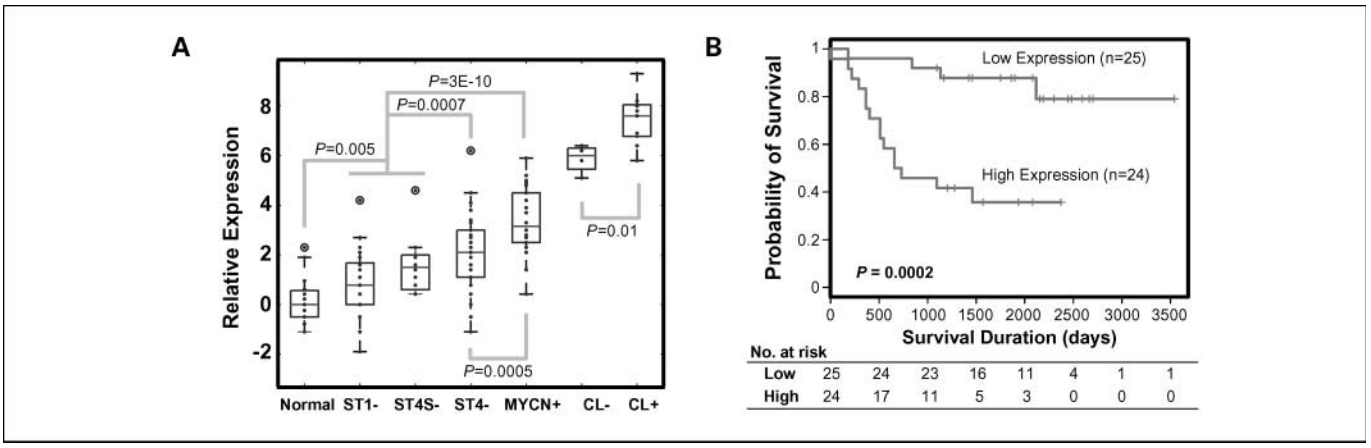

Fig. 4. High *MIRHG1* expression is significantly correlated with poor clinical outcome in neuroblastoma patients. *A,* box plot of *MIRHG1* expression in neuroblastoma primary tumors and cell lines shows that *MIRHG1* is expressed at a higher level in high-stage ($P = 0.0007$) and *MYCN*-amplified ($P = 0.0005$) neuroblastoma tumors as well as in the cell lines ($P = 0.01$). Normal samples express lower level of *MIRHG1* ($P = 0.005$). *Normal,* normal tissues; *ST,* stage; *CL,* cell line; +, *MYCN* amplification; -, *MYCN* not amplified. *B,* Kaplan-Meier curve for survival probability using the mRNA level of *MIRHG1* in a published NCI neuroblastoma patient cohort (23) shows that *MIRHG1* expression level can predict the outcomes of neuroblastoma patients. Median expression of *MIRHG1* was used as the cutoff, and *P* value is calculated using a log-rank test.

multiple human malignancies (15, 16). Among these 43 microRNAs, we found that 5 of 7 microRNAs in the OncomiR-1 cluster significantly correlated with the expression level of its host gene *MIRGH1* ($r = 0.37$-0.49; $P < 0.01$; Fig. 3), indicating that the level of *MIRGH1* transcript directly affected the level of these oncogenic microRNAs.

High MIRGH1 expression level is associated with aggressive behavior and poor outcome in neuroblastoma patients. Because OncomiR-1 has been reported to be directly downstream of MYC (17–20), and MYCN is often highly expressed in pediatric cancers including neuroblastoma and rhabdomyosarcoma (21, 22), we explored if the expression of OncomiR-1 host gene *MIRHG1* was associated with more aggressive phenotype in our neuroblastoma tumor gene expression database (23, 24).[6] We found a significantly higher expression level in *MYCN*-amplified tumors and cell lines as well as in higher stage (stage IV) compared with lower stage and more benign tumors (stage I and IVS; Fig. 4A). The highest expression was in cell lines particularly if the *MYCN* was amplified and the lowest expression was in normal human tissues ($n = 19$; Fig. 4A). Furthermore, we found that high expression of *MIRHG1* is significantly associated with poor outcome of neuroblastoma patients in our published cohort consisting of patients of all major stages with or without *MYCN* amplification (NCI cohort; $P = 0.0002$, log-rank test, or $P = 0.0011$, univariate Cox model; Fig. 4B and Table 2; ref. 23). A multivariate analysis showed that the prognostic power of *MIRHG1* expression is independent of current Children's Oncology Group risk-stratification ($P < 0.05$; Table 2).

Finally, we examined if our findings could be validated in another independent neuroblastoma patient cohort reported by Asgharzadeh et al. (25), which only included neuroblastoma patients with stage IV diseases without *MYCN* amplification (Children's Hospital Los Angeles cohort). Indeed, we observed higher *MIRHG1* expression in the patients with poor outcome in this data set ($P = 0.0008$; Fig. 5A), confirming that high expression level of *MIRHG1* was significantly associated with poor prognosis ($P = 0.0149$; Fig. 5B). Furthermore, we examined if the expression level of *MIRHG1* added any predictive power to the current Children's Oncology Group risk-stratification in this data set. Because the intermediate-risk patients in Children's Hospital Los Angeles cohort all survived, we could not build a Cox regression model using the Children's Oncology Group risk-stratification criteria in the multivariate analysis. Instead, we used the available risk factors (age and histology) in this analysis (Table 3). Multivariate analysis in this data set again showed that *MIRHG1* expression is a significant prognostic

Table 2. Survival analyses of NCI cohort

Parameter	HR	95% CI low	95% CI high	*P*
Univariate analysis				
Age (≥18 mo vs <18 mo)	14.9	4.28	52.2	2.2E-5
INSS stage (III&IV vs I&II)	13.7	1.83	103	0.011
MYCN (amp. vs non-amp.)	10.3	3.78	28.2	5.3E-6
COG risk (high vs low & intermediate)	30.1	4.00	227	9.4E-4
MIRHG1 (high vs low expression)	6.68	2.14	20.9	1.1E-3
Multivariate analysis				
COG risk (high vs low & intermediate)	22.9	3.00	175	2.5E-3
MIRHG1 (high vs low expression)	4.03	1.29	12.6	0.016

NOTE: Expression of *MIRHG1* predicts the outcome of neuroblastoma patients independently from current Children's Oncology Group risk-stratification in the NCI cohort (23). Cox regression models were used in both univariate and multivariate analyses.

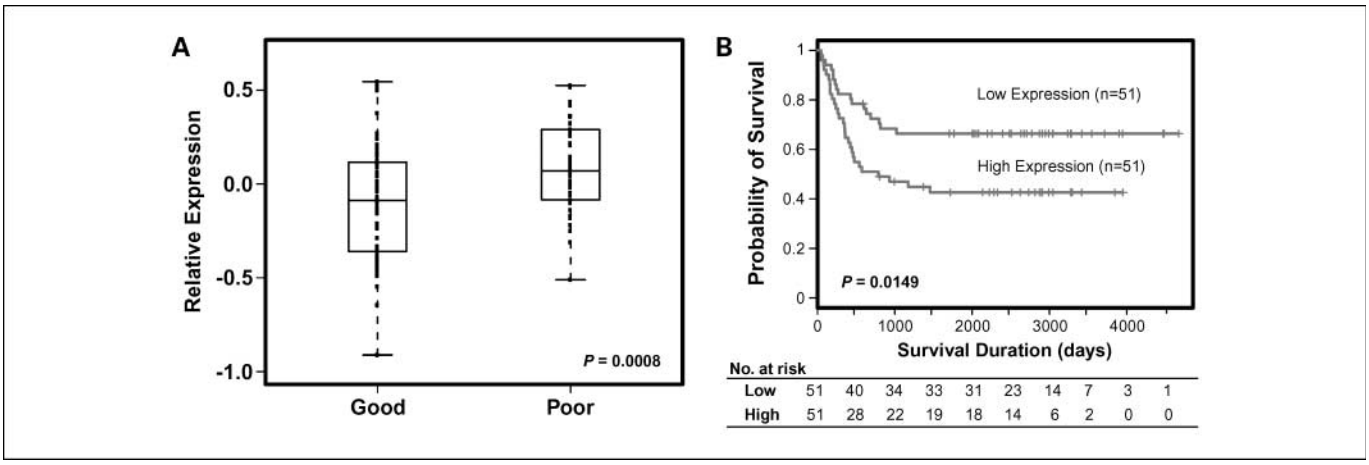

Fig. 5. Predictive power of *MIRHG1* expression is validated in an independent published Children's Hospital Los Angeles patient cohort, which only consists of neuroblastoma patients with stage IV diseases and without *MYCN* amplification (25). *A*, the high expression level of *MIRHG1* is also observed in the patients of poor outcome ($P = 0.0008$). The expression of *MIRHG1* significantly correlated with adverse outcome in a Kaplan-Meier plot ($P = 0.0149$; *B*) and Cox models ($P < 0.05$; Table 3). Median expression of *MIRHG1* was used as the cutoff.

marker independent from Children's Oncology Group risk-stratification factors (Table 3).

Discussion

The expression patterns of microRNA represent novel methods for categorizing pediatric cancers and have potential advantages over mRNA profiling. First, the relatively smaller number of microRNA in the human genome (~1,000 for microRNA versus tens of thousands for mRNA) makes their profiles less complex. Second, because of the short length of microRNAs, they are better preserved in clinical samples such as tumor specimen and paraffin sections (26, 27). Reports have suggested that microRNA expression profiles may be better predictors of diagnosis and clinical outcome of human diseases than mRNA-based methods (28, 29); however, a recent study on NCI-60 cell lines showed that mRNA profiles are more informative for discriminating tissue types than microRNA profiles (30). In our study, the 57 samples clustered according to the major cancer types using the microRNA expression profiles (Fig. 1). However, the microRNA expression profiles did not perform as well as mRNA expression profiles (7, 8) in separating these samples into their cancer types by hierarchical clustering (data not shown). Therefore, the value of using global microRNAs profiles in cancer classification is still unclear.

We identified 14 significantly differentially expressed microRNAs distinguishing neuroblastoma from rhabdomyosarcoma in these xenografts. Of these, 8 were validated by TaqMan reverse transcription-PCR in an independent cohort of primary human tumor samples, indicating the potential utility of these microRNAs as tumor-specific biomarkers for tumor classification. However, due to the rarity of pediatric solid tumors, this study focused on validation of microRNAs only in neuroblastoma and rhabdomyosarcoma. Although the xenografts and cell lines are kept in an artificial environment, studies have shown that they express a large panel of genes resembling their corresponding human tumors (7, 8). Furthermore, studies in the xenografts and cell lines have yielded valuable information in a preclinical setting such as the Pediatric Preclinical Testing Program[8] to identify effective agents for these tumors (7, 8). Therefore, we believe our approach in this study is valid and the results from this study showed that we identified cancer-specific microRNAs. However, these microRNA signatures need to be

[8] http://pptp.stjude.org/

Table 3. Survival analyses of CHLA cohort

Parameter	HR	95% CI low	95% CI high	*P*
Univariate analysis				
Age (≥18 mo vs <18 mo)	23.2	5.61	96.0	1.4E-5
Histology (unfavor vs favor)	19.0	4.60	78.6	4.8E-5
MIRHG1 (high vs low expression)	2.07	1.14	3.78	0.017
Multivariate analysis				
Age (≥18 mo vs <18 mo)	21.9	1.05	455	0.046
Histology (unfavor vs favor)	1.03	0.05	21.6	0.980
MIRHG1 (high vs low expression)	1.87	1.02	3.42	0.043

NOTE: The expression of *MIRHG1* significantly correlated with adverse outcome in Cox models ($P < 0.05$). Median expression of *MIRHG1* was used as the cutoff.

validated in a much larger study with human tumors before clinical use as diagnostic markers.

In addition to potential use a diagnostic biomarkers, these differentially expressed microRNAs may shed light to our understanding of the biology of these pediatric tumors, because microRNA expression has been reported to associate with tissue differentiation (31). Thus, tumor-specific microRNA profiles may reveal not only the tissue origin of cancers but also its biology in tumorigenesis. For example, miR-133a is expressed during normal muscle development (31) and was found by us to be expressed abundantly in rhabdomyosarcoma. Recently miR-133a has been shown to play a critical role in the regulation of myocyte growth (32, 33). Similarly, we found that miR-7, 124a, 137, and 218 were expressed at a high level in neuroblastoma samples, and their expression has been reported to be specific in neural tissues during zebrafish development (34). The correlation between the expression levels of miR-218 and its host gene, *SLIT3*, is the highest in our analysis (Supplementary Table S2). *SLIT* family members have been implicated to play a critical role in the formation of central nervous system (35). This indicates that the expression of miR-218 from the *SLIT3* transcript may also play a role in the differentiation of neural tissues. Intriguingly, we have reported previously that *SLIT3* is overexpressed in the poor-prognosis neuroblastoma (23). The high correlation between *SLIT3* and miR-218 expression suggests that expression of miR-218 may also predict poor prognosis. Recently, Makeyev et al. have reported that miR-124 promotes neuronal differentiation through inducing nervous system–specific alternative splicing (36) showing the importance of this microRNA during neural development. Therefore, differentially expressed microRNAs are likely to play important roles in the normal tissue development as well as in the tumorigenesis of pediatric cancers.

Despite the increasing knowledge of microRNA expression patterns in different biological systems including cancers, the regulation of microRNA expression is largely unknown. We attempted to determine if the genomic location of microRNAs in relationship to their host genes affected the expression of microRNAs. Using this pediatric tumor data set containing both gene and microRNA expression profiles and a low stringent cutoff ($r > 0.34$), surprisingly we observed that only 63% of microRNAs residing within host genes in the same orientation showed expression patterns that correlated with their host genes (Fig. 3). Therefore, a gene unit is a more complicated functional transcription unit than the protein coding gene itself, indicating that there are multiple mechanisms to regulate microRNA levels other than sharing the common promoter with their host genes.

Among all the microRNAs and their host gene with highly correlated expression levels, OncomiR-1 (miR-17-92 cluster) and its nonprotein coding host gene, *MIRHG1* (also known as *C13orf25*), are of particular interest due to its oncogenic potential in human cancers (16, 17, 20). Although *MIRHG1* and the microRNAs in OncomiR-1 are correlated, the putative promoter of *MIRGH1* is ~2,000 bp upstream of the E-boxes at the OncomiR-1 locus (18, 20). It is therefore still possible that *MIRHG1* and OncomiR-1 are transcribed from different promoters but are still coregulated. OncomiR-1 has been shown to be directly transactivated by an important oncogene c-Myc (18), and *MYC* oncogene family members are important transcription factors often hyperactivated in many human cancers (37). Therefore, OncomiR-1 can mediate at least some of the oncogenic functions of MYC. Several recent studies have indicated that MYCN, another MYC family member, can up-regulate OncomiR-1 (19, 20, 38). Fontana et al. have shown that MYCN activates OncomiR-1 cluster by directly binding to its promoter (20). This is of particular interest because the *MYCN* gene is frequently amplified in neuroblastoma and rhabdomyosarcoma (21, 22), and this molecular characteristic is used in clinic to stratify treatment for patients with neuroblastoma. In this study, we have shown that the high expression of the OncomiR-1 host gene, *MIRHG1*, is correlated with tumors with not only *MYCN* amplification but also higher stages and poor prognosis. In addition, Fontana et al. have shown the tumorigenic role of OncomiR-1 cluster in neuroblastoma cells by promoting cell growth (20). Therefore, these studies indicate an important biological role of OncomiR-1 cluster in the aggressive form of neuroblastoma. These findings warrant future studies to characterize the oncogenic mechanisms of individual microRNA encoded in this cluster and explore the potential targeted therapies against these microRNAs.

In summary, we have shown that pediatric cancers cluster according to their diagnosis based on microRNA expression profiles, and we identified 8 tumor-specific microRNAs for rhabdomyosarcoma and neuroblastoma. In addition, we have shown evidence that the expression of microRNAs located within protein coding genes is coregulated with their host transcripts. Finally, we showed that the high expression of a microRNA cluster host gene *MIRHG1* is significantly associated with aggressive neuroblastoma. Our results indicate that *MIRHG1* may play an important biological role in aggressive neuroblastoma, and the predictive value of *MIRHG1* expression in neuroblastoma patients should be further validated in a much larger cohort in future studies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

This article is dedicated to our friend and colleague Dr. Stephen Qualman (Cooperative Human Tissue Network) who fought a valiant battle against pancreatic cancer.

We thank Dr. Ernest Kawasaki, David Petersen, Jonathon Paarlberg, and Temitope Fadiran for excellent technical support; Dr. Thomas Badgett for insightful comments on this article; and the Corporative Human Tissue Network for providing human tumor samples.

References

1. Bartel DP. microRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.
2. Lim LP, Lau NC, Garrett-Engele P, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 2005;433:769–73.
3. He L, Hannon GJ. microRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 2004;5:522–31.
4. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524–9.
5. Cimmino A, Calin GA, Fabbri M, et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005;102:13944–49.
6. Costinean S, Zanesi N, Pekarsky Y, et al. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proc Natl Acad Sci U S A 2006;103:7024–9.

7. Neale G, Su X, Morton CL, et al. Molecular characterization of the pediatric preclinical testing panel. Clin Cancer Res 2008;14:4572–83.

8. Whiteford CC, Bilke S, Greer BT, et al. Credentialing preclinical pediatric xenograft models using gene expression and tissue microarray analysis. Cancer Res 2007;67:32–40.

9. Khan J, Wei JS, Ringner M, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med 2001;7:673–9.

10. Wei JS, Song YK, Durinck S, et al. The MYCN oncogene is a direct target of miR-34a. Oncogene 2008;27:5204–13.

11. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006;34:D140–4.

12. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics 2004;20:1453–4.

13. Saldanha AJ. Java Treeview-extensible visualization of microarray data. Bioinformatics 2004; 20:3246–8.

14. Dabney AR. Classification of microarrays to nearest centroids. Bioinformatics 2005;21:4148–54.

15. Ota A, Tagawa H, Karnan S, et al. Identification and characterization of a novel gene, C13orf25, as a target for 13q31-32 amplification in malignant lymphoma. Cancer Res 2004;64:3087–95.

16. Hayashita Y, Osada H, Tatematsu Y, et al. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res 2005;65: 9628–32.

17. He L, Thomson JM, Hemann MT, et al. A microRNA polycistron as a potential human oncogene. Nature 2005;435:828–33.

18. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature 2005;435:839–43.

19. Schulte JH, Horn S, Otto T, et al. MYCN regulates oncogenic microRNAs in neuroblastoma. Int J Cancer 2008;122:699–704.

20. Fontana L, Fiori ME, Albini S, et al. Antagomir-17-5p abolishes the growth of therapy-resistant neuroblastoma through p21 and BIM. PLoS ONE 2008; 3:e2236.

21. Schwab M, Ellison J, Busch M, Rosenau W, Varmus HE, Bishop JM. Enhanced expression of the human gene N-myc consequent to amplification of DNA may contribute to malignant progression of neuroblastoma. Proc Natl Acad Sci U S A 1984;81: 4940–4.

22. Williamson D, Lu YJ, Gordon T, et al. Relationship between MYCN copy number and expression in rhabdomyosarcomas and correlation with adverse prognosis in the alveolar subtype. J Clin Oncol 2005;23:880–8.

23. Wei JS, Greer BT, Westermann F, et al. Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. Cancer Res 2004;64: 6883–91.

24. Krasnoselsky AL, Whiteford CC, Wei JS, et al. Altered expression of cell cycle genes distinguishes aggressive neuroblastoma. Oncogene 2005;24:1533–41.

25. Asgharzadeh S, Pique-Regi R, Sposto R, et al. Prognostic significance of gene expression profiles of metastatic neuroblastomas lacking MYCN gene amplification. J Natl Cancer Inst 2006;98: 1193–203.

26. Li J, Smyth P, Flavin R, et al. Comparison of miRNA expression patterns using total RNA extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells. BMC Biotechnol 2007;7:36.

27. Nelson PT, Baldwin DA, Scearce LM, Oberholtzer JC, Tobias JW, Mourelatos Z. Microarray-based, high-throughput gene expression profiling of microRNAs. Nat Methods 2004;1:155–61.

28. Lu J, Getz G, Miska EA, et al. microRNA expression profiles classify human cancers. Nature 2005;435:834–8.

29. Calin GA, Ferracin M, Cimmino A, et al. A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353:1793–801.

30. Blower PE, Verducci JS, Lin S, et al. microRNA expression profiles for the NCI-60 cancer cell panel. Mol Cancer Ther 2007;6:1483–91.

31. Wienholds E, Kloosterman WP, Miska E, et al. microRNA expression in zebrafish embryonic development. Science 2005;309:310–1.

32. Chen JF, Mandel EM, Thomson JM, et al. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet 2006;38:228–33.

33. Care A, Catalucci D, Felicetti F, et al. microRNA-133 controls cardiac hypertrophy. Nat Med 2007; 13:613–8.

34. Wienholds E, Plasterk RH. microRNA function in animal development. FEBS Lett 2005;579: 5911–22.

35. Itoh A, Miyabayashi T, Ohno M, Sakano S. Cloning and expressions of three mammalian homologues of *Drosophila* slit suggest possible roles for Slit in the formation and maintenance of the nervous system. Brain Res Mol Brain Res 1998;62:175–86.

36. Makeyev EV, Zhang J, Carrasco MA, Maniatis T. The microRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol Cell 2007;27: 435–48.

37. Nesbit CE, Tersak JM, Prochownik EV. MYC oncogenes and human neoplastic disease. Oncogene 1999;18:3004–16.

38. Chen Y, Stallings RL. Differential patterns of microRNA expression in neuroblastoma are correlated with prognosis, differentiation, and apoptosis. Cancer Res 2007;67:976–83.